Author's Accepted Manuscript

Unions in a Frictional Labor Market

Per Krusell, Leena Rudanko

PII: S0304-3932(16)30018-6
DOI: http://dx.doi.org/10.1016/j.jmoneco.2016.04.006
Reference: MONEC2851

To appear in: *Journal of Monetary Economics*

Received date: 27 March 2014
Revised date: 22 April 2016
Accepted date: 25 April 2016

Cite this article as: Per Krusell and Leena Rudanko, Unions in a Frictional Labo Market, *Journal of Monetary Economics* http://dx.doi.org/10.1016/j.jmoneco.2016.04.006

Unions in a Frictional Labor Market☆

Per Krusell

Institute for International Economic Studies, University of Gothenburg, NBER, and CEPR

Leena Rudanko*

Federal Reserve Bank of Philadelphia

Abstract

A labor market with search and matching frictions, where wage setting is controlled by a monopoly union that follows a norm of wage solidarity, is found vulnerable to substantial distortions associated with holdup. With full commitment to future wages, the union achieves efficient hiring in the long run, but hikes up wages in the short run to appropriate rents from firms. Without commitment, in a Markov-perfect equilibrium, hiring is too low both in the short and the long run. The quantitative impact is demonstrated in an extended model with partial union coverage and multiperiod union contracting.

Keywords: Labor unions, frictional labor markets, time inconsistency, limited commitment, long-term wage contracts

JEL: E02, E24, J31, J51, J64

1. Introduction

Labor unions play an important role in many labor markets in many countries. There is also a large body of literature within labor economics focusing on how union presence influences labor

☆We are grateful to Marina Azzimonti, Matteo Cacciatore, Steve Davis, Fatih Guvenen, William Hawkins, Patrick Kehoe, John Kennan, Guido Menzio, Fabien Postel-Vinay, Victor Rios-Rull, Robert Shimer, seminar and conference audiences, as well as the editor and referee for comments. Rudanko thanks the Hoover Institution for its hospitality and financial support, and the Fulbright Program for financial support. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

*Corresponding author. Research Department, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia PA 19106, United States. Tel. +1-215-574-3719, Fax. +1-215-574-4303.

Email address: leena.rudanko@gmail.com (Leena Rudanko)

market outcomes. Yet, there is relatively little work studying the impact of this institution on the labor market when this market is described as having frictions and featuring unemployment due to these frictions. Since search and matching models have come to play a central role as a workhorse for macroeconomic labor market analyses, this gap in the literature leaves open important questions: What is the impact of unions on unemployment and wages? How do unions affect how strongly unemployment varies over the business cycle? What institutional settings are desirable, when considering rules regarding union coverage?

The model can be interpreted as representing either the aggregate labor market or an industry labor market, but in either case, the focus is on the case of a "large" union, which has monopoly power over some group of workers. This case is particularly relevant for many European economies, in which there is a nationwide union or cooperation/agreements among unions representing different industries. It is also relevant in other settings in which workers cannot easily move across industries and competition among different unions within an industry is limited. The union is assumed to be fully rational, taking job creation into account when making its wage demands, and its objective to be the welfare of all workers covered by union wages.

In the model, all workers have the same productivity and fulfill equally productive jobs. We start with the view that union operations are governed by a norm of solidarity and egalitarianism among workers, which leads to the assumption that unions impose identical wages across these workers. This view can be motivated in part by the broad empirical evidence documenting that unions compress the distribution of wages. Such fairness is found to come at a nontrivial cost, however, as it leaves the unionized labor market vulnerable to a potentially severe holdup problem, which leads to inefficiently high wages and low job creation.

Under the egalitarian wage policy, the degree to which the union can commit to future wages becomes qualitatively and quantitatively important for outcomes.[1] If the union can fully commit to future wages, it attains an efficient level of unemployment in the long run. In the short run, how-

[1]The degree of commitment to wages is important in holdup problems in general, with full commitment potentially avoiding the holdup problem entirely. In the dynamic model with an egalitarian wage policy, the situation is more involved, however, because even in the union problem with full commitment there are some workers who were hired in the past and whose wages will in part be set after they have already been hired.

ever, unemployment is inefficiently high because the union uses its market power to raise current wages above the efficient level to extract rents from firms with preexisting matches. Specifically, labor market tightness is shown to be inefficiently low in the initial period but efficient from then on. These elements give rise to a time inconsistency: If a union had decided on a commitment plan yesterday, but had the opportunity to revise it today, the union would indeed revise the plan to benefit again from preexisting matches.

What would happen if the union did not have commitment to future wages? What effects would it have on the labor market? The paper answers this question by analyzing differentiable Markov perfect equilibria.[2] In a calibrated model, the presence of the union raises wages by 11%, consequently raising unemployment from 5% to 16%, and reducing output by 12%, relative to efficient outcomes. The distortions associated with the union diminish as the duration of union contracts increases, but this effect appears quantitatively weak; the effects remain very similar as duration varies from one to three years, viewed as the empirically relevant range of union contract durations (Taylor, 1983).

In a classic paper, Calmfors and Driffill (1988) reconsidered the impact of unions on the level of aggregate economic activity. It has long been recognized that unions, through their monopoly power in the labor market, tend to raise wages above their competitive levels, suggesting that a greater union presence in the labor market has a primarily negative impact on economic activity. Calmfors and Driffill (1988) propose an additional factor for understanding the cross-country evidence on unions: They argue that the degree of coordination in union bargaining works to counteract the negative effects of monopoly power. Our model generates a related hump-shaped relationship, illustrated in Section 4, which allows partial union coverage of the workforce. Because union wages tend to be higher than nonunion wages, greater union coverage tends to lead to higher unemployment in our model as well. But greater union coverage also increases the extent to which the union takes into account the effects of its wage demands on hiring, borne by union and nonunion workers alike, leading to moderation in union wage setting. As union coverage increases, the second effect eventually takes over the first, leading to a hump-shaped relationship.

[2]We extend the solution approach of Krusell et al. (2002) for the generalized Euler equation to allow solving for labor market outcomes with partial union coverage and multiperiod union contracting.

An important motivation for macroeconomists to consider unions has been the idea that union wages are less responsive to shocks, potentially helping to understand the observed variability of employment (see, e.g., Blanchard and Fischer, 1989, pp. 438–455). The model studied in Section 4 builds in significant stickiness in wages, because the union recontracts only every one to three years. The stickiness has a substantial impact on shock propagation in the model, with amplification in the responses of vacancy creation, employment, and output to shocks.

Finally, while we view egalitarianism as a characteristic of union operations, it is also shown that relaxing the egalitarian wage policy, for example by allowing a tenure premium in union wages, can provide the union sufficient instruments to avoid the holdup problem, perhaps entirely. In this case, the union extracts rents from firms with high wages for senior workers, while setting the wages of junior workers low enough to encourage hiring nevertheless. Unless the union runs into a binding constraint on how low the wages of junior workers can be (possibly negative), efficient hiring is attained. The model thus implies a rationale for a tenure premium in union wages.

Related literature. There are papers developing extensions of the Mortensen-Pissarides model with a union/unions governing wage determination. Perhaps closest in spirit to our paper is Pissarides (1986), which first introduces a monopoly union into the Pissarides (1985) framework and studies the impact on equilibrium outcomes in the labor market. As with the literature following it, Pissarides (1986) focuses on steady states, however, side-stepping the dynamic issues highlighted here.[3] The more recent papers are more applied: Garibaldi and Violante (2005) and Boeri and Burda (2009) study the effects of employment protection policies; Ebell and Haefke (2006) study the effects of product market regulation; and Acikgoz and Kaymak (2014) study the evolution of skill premia and unionization rates over time. These papers generally adopt frameworks

[3]Lockwood and Manning (1989) and Modesto and Thomas (2001) have studied union wage setting in labor markets in which firms face adjustment costs to labor, developing the idea that dynamic concerns become important for thinking about union decision-making when labor markets are not fully frictionless. The simple partial equilibrium quadratic adjustment cost framework adopted in these papers affords closed-form results that speak to the level of union wage demands, as well as to the speed of adjustment in firm-level employment. Our work brings these ideas into an equilibrium framework, which allows us to consider unemployment and vacancy creation as well.

imposing exogenous wage compression into union wage setting, with the exception of Taschereau-Dumouchel (2011), who develops a framework where it is endogenous. Delacroix (2006) extends the framework of Ebell and Haefke (2006) to capture the U-shaped relationship between the degree of coordination in union bargaining and economic performance postulated by Calmfors and Driffill (1988).

Other related work includes Acemoglu and Pischke (1999), who argue that union wage compression across workers of differing skill levels can encourage firms to provide training; Alvarez and Veracierto (2000), who study an extension of the Lucas and Prescott (1974) island model with unions quantitatively, considering several alternative ways of modeling union behavior (worker coalition vs. union boss, equal treatment vs. insider-outsider framework); and Alvarez and Shimer (2011), who study a further extension of the Lucas and Prescott model that allows search also on the islands, emphasizing the role of seniority for union hiring and layoff decisions.

The paper is organized as follows: Section 2 begins with a brief overview of the empirical evidence on unions. Section 3 analyzes the benchmark model: first, a one-period model to provide intuition, and then an infinite-horizon model with and without commitment. Section 4 turns to a quantitative illustration in the context of an extended model, and Section 5 concludes.

2. Evidence on unions, wages and unemployment

Most workers in the OECD, outside the U.S., have their wages determined by union agreements. This cross-country evidence is discussed by Nickell and Layard (1999), who report that in most European countries, the share of workers covered by union wages exceeds 70%. An important feature of the cross-country evidence is that union coverage rates—the share of the labor force whose wages are determined by union wage bargaining—generally exceed union membership rates outside the US. Even in countries in which union membership rates are low, such as France, within firms many nonunion workers are paid the union wage, and in many countries, union wages are legally extended to cover nonunion firms as well. Visser (2003) also documents union membership and coverage rates across countries, reporting an average coverage rate of 73% across European countries for the period 1985–1997. While union membership has been on the

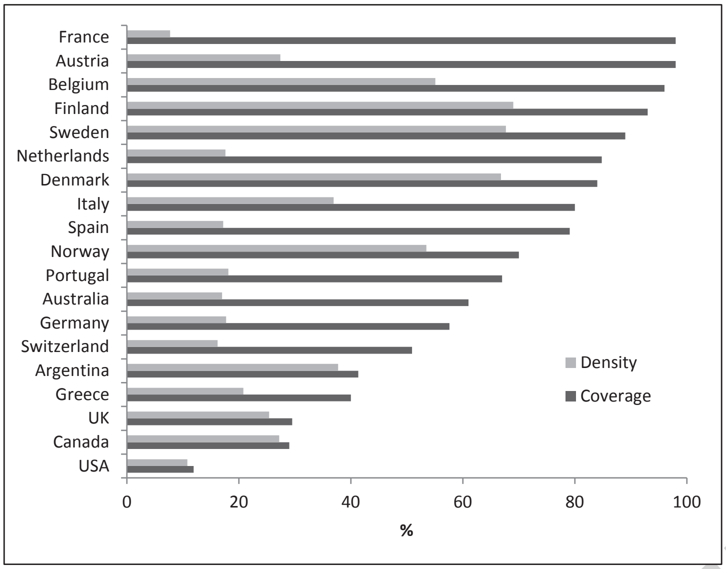


Figure 1: Unionization Rates across Countries in 2013

Notes: The figure displays union coverage and density rates for 2013 based on data available from the International Labor Organization (`www.ilo.org/ilostat/`). Coverage refers to the share of workers to whom a collective agreement applies, and density to the share who are union members. The density for Argentina is from 2008, the most recent figure reported.

decline in Europe as well as in the U.S., coverage levels remain substantially higher in Europe, as Figure 1 shows.[4]

In terms of the effects of unions, Nickell and Layard (1999) show that a cross-country regression of unemployment on measures of union membership and coverage reveals a positive relationship between union presence and unemployment. But there is also significant heterogeneity across countries in the degree of centralization and coordination in union bargaining, as highlighted by Calmfors and Driffill (1988), and it turns out that this positive relationship between union presence and unemployment can be partly offset by measures of coordination in bargaining.

Nickell and Layard (1999) also report that union membership is associated with higher wages on the individual level across countries. An extensive literature has studied this union/nonunion wage gap, using a variety of data sources and econometric approaches. Lewis (1986) reviews the literature for the U.S., concluding that the evidence points to an upper bound of 15% for the

[4]See also Visser et al. (2015) for more background and broader cross-country evidence.

union wage gap. More recently, Blanchflower and Bryson (2003) confirm that the estimates of the wage gap have remained relatively stable, with perhaps a modest decline over time. They also report estimates across countries, noting that in many European countries the extensive coverage of union wages reduces these gaps. An important concern with the estimates of the union wage gap in general involves selection on unobservables: It is likely that higher union wages attract better workers, but the data do not allow these differences to be controlled for properly, biasing the estimates of the wage gap. When DiNardo and Lee (2004) adopt a regression discontinuity design to get around some of the issues, they find a negligible wage gap, seemingly contradicting a large body of evidence.[5]

A robust finding appears to be that unions reduce wage inequality, compressing the distribution of wages (Card et al., 2003). Do they compress wages across degrees of seniority as well? Certainly formal pay scales appear to be common in union compensation practices, but arguably wages rise with tenure in nonunion settings as well. Perhaps because unions tend to compress the distribution of wages, a number of earlier studies have actually reported a stronger association between tenure and earnings in nonunion settings. But properly estimating returns to tenure is challenging and the comparison is confounded by the fact that the estimates tend to be biased by worker and job heterogeneity, generally found to be greater in nonunion than in union settings.[6] Recognizing these challenges, Abraham and Farber (1988) find a stronger association between tenure and earnings in the unionized setting, supporting the idea that seniority plays an important role in union operations. At the same time, Topel (1991) finds no significant difference in returns to tenure based on union status. Again, data limitations leave us short of a conclusive answer, but the evidence in favor of overall wage compression does appear to be robust.

[5]Their study focuses on close union election outcomes in the U.S. Of course, it is possible that wage gaps in workplaces with close election outcomes are smaller than in those with clear-cut outcomes, and that wages in newly unionized workplaces are different from those with an established union presence.

[6]The magnitude of returns to tenure is a debated topic; see, for example, Altonji and Williams (2005) and Buchinsky et al. (2010).

3. The model

This section begins with a description of the simple Mortensen-Pissarides search and matching environment that the analysis is based on. A monopoly union is then introduced, and its behavior characterized, within that framework.

A frictional labor market. Time is discrete and the horizon infinite. The economy is populated by a continuum of measure one identical workers, together with a continuum of identical capitalists who employ these workers. All agents have linear utility and discount the future at rate $\beta < 1$. Capitalists have access to a linear production technology, producing z units of output per period for each worker employed. In addition to this market production technology, unemployed workers also have access to a home production technology, producing $b(< z)$ units of output per period.

The labor market is frictional, requiring capitalists seeking to hire workers to post vacancies. The measure of matches in the beginning of the period is denoted by $n \in [0, 1]$, leaving $1 - n$ workers searching for jobs. Searching workers and posted vacancies are matched according to a constant-returns-to-scale matching function $m(v, 1-n)$, where v is the measure of vacancies. With this, the probability with which a searching worker finds a job within a period can be written $\mu(\theta) = m(\theta, 1)$, and the probability with which a vacancy is filled $q(\theta) = m(1, 1/\theta)$, where $\theta = v/(1-n)$ is the labor market tightness. It is assumed that $\mu'(\theta)$ is positive and decreasing and $q'(\theta)$ negative and increasing. With this, employment equals n plus the measure of new matches, $\mu(\theta)(1-n)$. Jobs are destroyed each period with probability δ. Thus, the measure of matches evolves over time according to the law of motion

$$n_{t+1} = (1-\delta)\underbrace{(n_t + \mu(\theta_t)(1-n_t))}_{\text{employed}_t}. \tag{1}$$

Firms. Capitalists operate production through firms, and these firms need to post vacancies to find workers, at a cost κ per vacancy. Competition drives profits from vacancy creation to zero, with firms taking into account the union wage-setting behavior today and in the future. The zero-profit condition thus determines the current market tightness according to current and future wages

$\{w_{t+s}\}_{s=0}^{\infty}$ as follows:

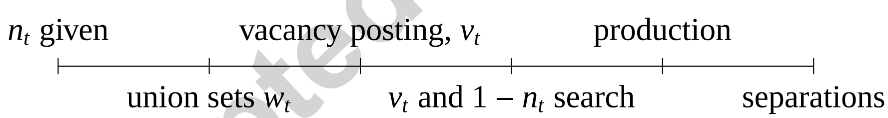
$$\kappa = q(\theta_t) \sum_{s=0}^{\infty} \beta^s (1-\delta)^s [z - w_{t+s}]. \tag{2}$$

Union. Wages are set unilaterally by a labor union, with universal coverage. The union sets wages to maximize the welfare of all workers, with equal pay for all those employed.[7] The union objective thus becomes

$$\sum_{t=0}^{\infty} \beta^t [\underbrace{(n_t + \mu(\theta_t)(1-n_t))}_{\text{employed}_t} w_t + \underbrace{(1-n_t)(1-\mu(\theta_t))}_{\text{unemployed}_t} b]. \tag{3}$$

The union takes as given the evolution of employment according to equation (1). It also internalizes the effect of its wage-setting decisions on hiring. Therefore, the union's problem is to choose a sequence of wages $\{w_t\}_{t=0}^{\infty}$ to maximize the objective (3) subject to the law of motion (1) and zero-profit condition (2). The union must also respect the constraint that the firms, at each point in time, make a nonnegative present value of profits on existing matches, as they could simply end them otherwise. This is implied by positive vacancy posting, however, because if firms posting vacancies break even, existing matches must have strictly positive value.

The following timeline summarizes the events in period t



Given the path of wages $\{w_t\}_{t=0}^{\infty}$, then, equation (2) determines the path of market tightness $\{\theta_t\}_{t=0}^{\infty}$, which in turn determines the evolution of employment.

3.1. One-period example

To illustrate key forces at play, the impact of the union is first considered in a very simple setting: a one-period version of the previous economy. Many features present here will be present in the subsequent analysis.

[7] Note that if one normalizes $b = 0$, then the union objective becomes the total wage bill.

Planner. A natural starting point is the efficient benchmark—the output-maximizing level of vacancy creation a social planner would choose. Here the planner solves the problem

$$\max_{\theta} \big(\underbrace{n + \mu(\theta)(1-n)}_{\text{employed}}\big) z + \underbrace{(1-n)(1-\mu(\theta))}_{\text{unemployed}} b - \underbrace{\theta(1-n)}_{\text{vacancies}} \kappa, \tag{4}$$

taking as given preexisting matches n. The planner's optimum is characterized by the first-order condition $-\kappa + \mu'(\theta)(z-b) = 0$, which pins down θ independent of n. For concreteness, consider the matching function $m(v,u) = vu/(v+u)$, such that $\mu(\theta) = \theta/(1+\theta)$. In this case, the planner's optimum is given by $\theta^p = \sqrt{(z-b)/\kappa} - 1$, with market tightness an increasing function of market productivity. Of course, it must be that $z - b > \kappa$ for vacancy creation to be optimal.

Union. The union instead aims to maximize the welfare of workers

$$\big(\underbrace{n + \mu(\theta)(1-n)}_{\text{employed}}\big) w + \underbrace{(1-n)(1-\mu(\theta))}_{\text{unemployed}} b, \tag{5}$$

by choice of w and θ, subject to the zero-profit condition: $\kappa = q(\theta)(z-w)$. The tradeoff the union faces here is that while higher wages increase the welfare of employed workers, they also reduce the job-finding probability because of reduced job creation.

To see how this problem relates to the planner's problem, one can use the zero-profit condition to solve for the wage, as $w = z - \kappa/q(\theta)$, and substitute it into the union objective to yield a maximization problem in θ only:

$$\max_{\theta} \big(\underbrace{n + \mu(\theta)(1-n)}_{\text{employed}}\big)\Big(z - \frac{\kappa}{q(\theta)}\Big) + \underbrace{(1-n)(1-\mu(\theta))}_{\text{unemployed}} b \tag{6}$$

$$= \max_{\theta} - \underbrace{\frac{n\kappa}{q(\theta)}}_{\text{capitalists' share}} + \underbrace{(n + \mu(\theta)(1-n))z + (1-n)(1-\mu(\theta))b - \theta(1-n)\kappa}_{\text{planner's objective}}, \tag{7}$$

also taking as given n.[8] From the second line, one can see that the union objective differs from the planner's objective only by the term $-\frac{n\kappa}{q(\theta)}$. To understand how the two objectives relate to each

[8]This substitution assumes some vacancy creation is optimal. The union could also opt to simply set $w = z$ in the original problem, achieving the value $b + n(z-b)$ for the objective (forgoing vacancy costs entirely). To ensure the solution in the text is optimal, it is necessary to make sure the value of the objective exceeds this value.

other, recall that while the planner cares about all agents in the economy, the union only cares about workers. The union objective thus equals the planner's objective less the capitalists' share of total output: the profits on existing matches $n(z-w) = \frac{n\kappa}{q(\theta)}$, where the equality follows from the zero-profit condition.

An interior union optimum is characterized by the first-order condition $-\kappa+\kappa\frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}+\mu'(\theta)(z-b) = 0$, which implies that the union's choice of θ does depend on n. In our example, an interior union optimum is given by $\theta = \sqrt{1-n}\sqrt{(z-b)/\kappa} - 1$. Labor-market tightness is thus again an increasing function of market productivity but now decreases in preexisting matches. Clearly, the union implements the socially optimal level of vacancy creation if $n = 0$. But if $n > 0$, the union has an incentive to raise wages above the efficient level, to appropriate surpluses from firms with existing matches.[9]

Finally, note that a nonegalitarian union would instead solve the problem

$$\max_{\theta,w^e,w^n} nw^e + \mu(\theta)(1-n)w^n + (1-\mu(\theta))(1-n)b \tag{8}$$

$$\text{s.t. } q(\theta)(z-w^n) = \kappa, \tag{9}$$

$$w^e \leq z, \tag{10}$$

where the union is allowed to pay different wages to newly hired workers, w^n, and workers in existing matches, w^e. Allowing different wages for the two groups immediately implies that the union sets $w^e = z$. Substituting this into the union objective then yields the planner objective above, along with the same condition for optimal hiring: $-\kappa+\mu'(\theta)(z-b) = 0$. With this market tightness, the wage in new matches is then given by $w^n = z - \kappa/q(\theta)$, implying a tenure premium in union wages: $w^n < w^e$.

This non-egalitarian case demonstrates that the inefficiency in the initial union problem stems

[9]Introducing curvature into the problem via a concave production function or a convex vacancy cost would bring about an added distortion reminiscent of that in static union problems, which also works to reduce hiring below efficient. The linearity of the baseline Mortensen-Pissarides model thus serves to isolate the dynamic distortion emphasized in this paper from the distortion appearing in static analyses of unionized labor markets. It also allows a relatively straightforward comparison of dynamics between the unionized labor market and efficiency, by making the efficient dynamics simple to characterize. See the online appendix, Section B, for more.

from the constraint to treat workers identically.[10] The theory thus implies a rationale for tenure premia in union wages, which could—in the absence of a binding lower bound on the wages of junior workers—even allow the union to attain efficient hiring.

The next section returns to the dynamic infinite horizon setting, where the measure of initial matches is endogenous.

3.2. *Efficient outcomes*

Beginning with the efficient outcome provides a useful benchmark for characterizing union wage-setting also when the time horizon is infinite. The planner now chooses a sequence $\{\theta_t\}_{t=0}^{\infty}$, with $\theta_t \geq 0$, to maximize

$$\sum_{t=0}^{\infty} \beta^t [\underbrace{(n_t + \mu(\theta_t)(1-n_t))}_{\text{employed}_t} z + \underbrace{(1-n_t)(1-\mu(\theta_t))}_{\text{unemployed}_t} b - \underbrace{\theta_t(1-n_t)}_{\text{vacancies}_t} \kappa] \tag{11}$$

$$\text{s.t. } n_{t+1} = (1-\delta) \underbrace{(n_t + \mu(\theta_t)(1-n_t))}_{\text{employed}_t}, \tag{12}$$

with n_0 given.

For what comes later, it will be useful to formulate problems recursively. The recursive form for the planner's problem reads

$$V^p(n) = \max_{\theta} (n + \mu(\theta)(1-n))z + (1-n)(1-\mu(\theta))b - \theta(1-n)\kappa + \beta V^p(N(n,\theta)), \tag{13}$$

where $N(n,\theta) \equiv (1-\delta)(n+\mu(\theta)(1-n))$. Notice that the state variable is n, the number of matches at the beginning of the period, and that the control variable—market tightness θ—determines n' according to the law of motion $N(n,\theta)$.

The first-order condition, assuming an interior solution, is

$$\kappa = \mu'(\theta)(z - b + \beta(1-\delta)V^{p\prime}(n')). \tag{14}$$

[10]These distortions arise because search frictions render existing matches a form of firm-specific capital, which is subject to a holdup problem. As is typically the case, the degree of commitment to wages is important for the severity of the holdup problem. In the extreme case, if wages are set after vacancy creation takes place (rather than before), the union would simply set (both) wages equal to z, with no new hiring taking place. The timing here allows the union to commit to wages before vacancy creation, however, making outcomes less severe.

It equalizes the cost of an additional vacancy, κ, to its benefits: an increase in matches of $\mu'(\theta)$, with each new worker delivering the flow surplus $z - b$ today, together with a continuation value reflecting future flow surpluses.

The envelope condition gives the value of an additional beginning-of-period match, as

$$V^{p\prime}(n) = (1 - \mu(\theta) + \theta\mu'(\theta))(z - b + \beta(1 - \delta)V^{p\prime}(n')). \tag{15}$$

This value takes into account that the increase in initial matches hampers current hiring by shrinking the pool of searching workers. To see this in the expression, note that the derivative of the matching function with respect to unemployment, $m_u(\theta, 1)$, equals $\mu(\theta) - \theta\mu'(\theta)$.

Eliminating the derivative of the value function in (14) yields the Euler equation

$$\frac{\kappa}{\mu'(\theta)} = z - b + \beta(1 - \delta)(1 - \mu(\theta') + \theta'\mu'(\theta'))\frac{\kappa}{\mu'(\theta')}. \tag{16}$$

This equation states the efficiency condition for the Mortensen-Pissarides model, solving a tradeoff between the costs and benefits of creating a new match today. The cost of an additional match today is $\kappa/\mu'(\theta)$: the cost of a vacancy, κ, times the measure of vacancies required for one match.[11] The benefits of an additional match include the flow surplus $z-b$ today, together with the expected value of the match next period. The expected value takes into account that the match survives to the next period with probability $1 - \delta$, and that the increase in matches shrinks the pool of searching workers tomorrow, so that any planned vacancy creation next period will yield fewer matches, leading to a net increase in matches of $1 - \mu(\theta') + \theta'\mu'(\theta')$. Finally, the value of a match tomorrow is again given by $\kappa/\mu'(\theta')$.

Note that the planner's Euler equation does not feature the state variable n explicitly at all, so a natural guess for the solution is a constant tightness independent of n. It is straightforward to show that the planner's value function is linear in n, and the efficient allocation is thus characterized by a constant market tightness $\theta_t = \theta^p$, for all $t \geq 0$.

[11]Since a unit increase in vacancies increases market tightness by $1/(1 - n)$ units, and a unit increase in market tightness yields $(1 - n)\mu'(\theta)$ new matches, one new vacancy creates $\mu'(\theta)$ new matches.

3.3. *A union with commitment*

Turning to the unionized labor market, consider the problem of the egalitarian union choosing a sequence of wages $\{w_t\}_{t=0}^{\infty}$ to maximize the objective (3) subject to the law of motion (1) and zero profit condition (2) holding at each point in time.

To relate the union problem to the planner's problem, one can again use the zero-profit conditions to rewrite the union objective. To this end, note that the union's choice of a sequence of wages determines, at each instant, the expected present value of union wages paid out over the course of an employment relationship: $W_t = \sum_{s=0}^{\infty} \beta^s (1-\delta)^s w_{t+s}$. The sequence $\{W_t\}_{t=0}^{\infty}$ further pins down the sequence $\{\theta_t\}_{t=0}^{\infty}$ through the zero-profit conditions, assuming some vacancy creation occurs each period. Conversely, given a sequence $\{\theta_t\}_{t=0}^{\infty}$, one can back out per-period wages by first using the zero-profit condition to find W_t each period, and then computing wages as $w_t = W_t - \beta(1-\delta)W_{t+1}$.

Using the zero-profit condition to eliminate wages, the union objective (3) can be written as:

$$-\frac{n_0 \kappa}{q(\theta_0)} + \sum_{t=0}^{\infty} \beta^t [(n_t + \mu(\theta_t)(1-n_t))z + (1-n_t)(1-\mu(\theta_t))b - \theta_t(1-n_t)\kappa], \tag{17}$$

revealing an identical objective to that of the planner except for the first term.[12] This term—familiar from the one-period example—reflects the share of the present discounted value of output accruing to capitalists. To see this, note that the capitalists' share, i.e., the present value of profits to firms, can be written as

$$n_0 \sum_{t=0}^{\infty} \beta^t (1-\delta)^t [z - w_t] + \sum_{t=0}^{\infty} \beta^t [\mu(\theta_t)(1-n_t) \sum_{s=0}^{\infty} \beta^s (1-\delta)^s [z - w_{t+s}] - \theta_t(1-n_t)\kappa]. \tag{18}$$

Here, the first term captures the present value of profits on existing matches, and the second those on new vacancies created in periods $t = 0, 1, \ldots$. The expression reduces to representing initial matches only, however, as free entry drives the present value of profits to new vacancies to zero.[13] Preexisting matches, on the other hand, are due a strictly positive present value of profits, because these firms paid the vacancy cost in the past, anticipating positive profits in the future to make up for it. Using the zero-profit condition, this remaining present value can be expressed as $n_0\kappa/q(\theta_0)$.

[12] See online appendix, Section A.

[13] The second term in equation (18) can be written as $\sum_{t=0}^{\infty} \beta^t (1-n_t)\theta_t [q(\theta_t) \sum_{s=0}^{\infty} \beta^s (1-\delta)^s [z - w_{t+s}] - \kappa]$, which equals zero because of the free entry condition (2).

The union objective (17) reflects the fact that while the planner maximizes the present discounted value of output, the union only cares about the workers' share of it. As a result, the union will have an incentive to appropriate some of this present value from capitalists by raising wages above the efficient level—and this is exactly how the solutions to the two problems will differ.

Proposition 1. *If the union is able to commit to future wages, hiring is efficient after the initial period. In the initial period, hiring is efficient if* $n_0 = 0$ *and below efficient if* $n_0 > 0$.

Note that after the initial period, the union effectively solves the planner's problem (13), and consequently chooses the planner's solution $\theta_t = \theta^p \; \forall t \geq 1$. In the initial period, however, the union chooses θ_0 to maximize

$$-\frac{n_0\kappa}{q(\theta_0)} + (n_0 + \mu(\theta_0)(1-n_0))z + (1-n_0)(1-\mu(\theta_0))b - \theta_0(1-n_0)\kappa + \beta V^p(N(n_0,\theta_0)), \tag{19}$$

where n_0 is given, and V^p solves the planner's problem (13).[14]

Deriving the optimality condition for this initial period is straightforward, using the same methods as above. Using the fact that the efficient market tightness θ^p will prevail in subsequent periods, the resulting condition can be written as

$$[1 - \frac{n_0}{1-n_0}\frac{q'(\theta_0)}{q(\theta_0)^2}]\frac{\kappa}{\mu'(\theta_0)} = z - b + \beta(1-\delta)(1-\mu(\theta^p) + \theta^p\mu'(\theta^p))\frac{\kappa}{\mu'(\theta^p)}. \tag{20}$$

Comparing with the efficiency condition (16), the cost of creating an additional match today (on the left) is higher for the union than for the planner. This occurs because in order to increase hiring, the union must lower wages, giving up some of the surplus it could have appropriated from firms with existing matches. Moreover, the more existing matches there are, the greater this additional cost.

Using the efficiency condition (16), equation (20) can be further rewritten as

$$[1 - \frac{n_0}{1-n_0}\frac{q'(\theta_0)}{q(\theta_0)^2}]\frac{1}{\mu'(\theta_0)} = \frac{1}{\mu'(\theta^p)}. \tag{21}$$

[14] Again, using the zero-profit condition to substitute out wages assumes positive vacancy creation each period. The union could, as an alternative, also choose to set the initial present value of wages so high as to shut down hiring in the first period entirely, allowing matches to depreciate. This becomes more attractive when initial matches are plentiful.

Because $q'(\theta) < 0$ and $\mu'(\theta)$ is decreasing, this equation implies that the market tightness will generally be lower in the initial period than the efficient value it takes on after that, and the more initial matches, the lower its initial value. Thus, as in the one-period example, the initial market tightness depends negatively on the measure of existing matches. This is a key feature of the model, which becomes even more important when the union does not have commitment.

That the outcome in the initial period differs from later periods reflects a time inconsistency issue in the union wage-setting problem. If the union were to reoptimize after the initial period, it would face a different objective and choose a different path of wages. While the union can thus get relatively close to the efficient outcome when it can commit, this immediate time inconsistency begs the question: What happens if the union cannot commit to future actions? To study time-consistent union decision-making, the next section turns to a game-theoretic setting, which will be based on the recursive formulation of the union problem set up above.

3.4. A union without commitment

The union problem (19) suggests that if the union were to reoptimize at any date, its choice of initial θ would depend on n, the measure of matches in the beginning of the period. In particular, a higher n should imply a lower θ. How would outcomes change if the union could not commit to not reoptimizing? In answering this question, this paper focuses on (differentiable) Markov perfect equilibria with n as a state variable. That n is a payoff- and action-relevant state variable should be clear from the problem under commitment. In a Markov perfect equilibrium, the union anticipates its future choices of θ to depend (negatively) on n, a relationship labeled $\Theta(n)$. The task, then, is to characterize $\Theta(n)$.

The function $\Theta(n)$ solves a problem similar to (19), namely

$$\Theta(n) \equiv \arg\max_{\theta} -\frac{n\kappa}{q(\theta)} + (n + \mu(\theta)(1-n))z + (1-n)(1-\mu(\theta))b - \theta(1-n)\kappa + \beta V(N(n,\theta)), \quad (22)$$

where the continuation value V satisfies the recursive equation

$$V(n) = (n + \mu(\Theta(n))(1-n))z + (1-n)(1-\mu(\Theta(n)))b - \Theta(n)(1-n)\kappa + \beta V(N(n,\Theta(n))). \quad (23)$$

Here, the union recognizes that its future actions will follow $\Theta(n)$, and this is reflected in the continuation value $V(n)$. Because $\Theta(n)$ will generally not be efficient, $V(n)$ will not equal $V^p(n)$, the continuation value under commitment.

A *Markov perfect equilibrium* is defined as a pair of functions $\Theta(n)$ and $V(n)$ solving (22)–(23) for all n. The functions are assumed to be differentiable, and equilibria characterized based on this assumption.

From equation (22), the first-order condition for market tightness becomes

$$[1 - \frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}]\kappa = \mu'(\theta)(z - b + \beta(1-\delta)V'(n')), \tag{24}$$

and the equation paralleling the envelope condition—now not formally an envelope condition since the union does not agree with its future decisions—becomes

$$\begin{aligned} V'(n) =& (1 - \mu(\theta) + \theta\mu'(\theta))(z - b + \beta(1-\delta)V'(n')) \\ &+ \mu'(\theta)(\Theta'(n)(1-n) - \theta)(-\frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}\frac{\kappa}{\mu'(\theta)}). \end{aligned} \tag{25}$$

Equation (25) is derived by differentiating equation (23) and using equation (24) to arrive at a formulation close to the equivalent condition (15) for the planner. Compared with the planner's envelope condition, this equation includes some additional terms, which appear because the envelope theorem does not hold. These terms work to reduce the value of additional initial matches n, as the union sets the market tightness too low—following $\Theta(n)$—and to an extent that increases in n.

One can further combine the above two equations to eliminate V', obtaining

$$\begin{aligned} \underbrace{[1 - \frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}]\frac{\kappa}{\mu'(\theta)}}_{\text{cost of match today}} =& z - b + \beta(1-\delta)[(1 - \mu(\theta') + \theta'\mu'(\theta'))\underbrace{[1 - \frac{n'}{1-n'}\frac{q'(\theta')}{q(\theta')^2}]\frac{\kappa}{\mu'(\theta')}}_{\text{value of match tomorrow}} \\ &\underbrace{+\mu'(\theta')(\Theta'(n')(1-n') - \theta')(-\frac{n'}{1-n'}\frac{q'(\theta')}{q(\theta')^2})\frac{\kappa}{\mu'(\theta')}}_{\text{loss in value from lack of commitment}}], \end{aligned} \tag{26}$$

which is a *generalized Euler equation*. It is a functional equation in the unknown policy function Θ, where the derivative of Θ appears. The equation is written in a short-hand way: θ is short for $\Theta(n)$, θ' is short for $\Theta(N(n, \Theta(n)))$, and n' is short for $N(n, \Theta(n))$. The task is to find a function Θ that solves this equation for all n. Note that in contrast to the planner's Euler equation, n appears nontrivially in this equation and will generally matter for the tightness. It is easily verified that a constant Θ will not solve the equation.

Equation (26), as with the planner's Euler equation (16), represents the tradeoff between the costs and benefits of creating matches today. The cost of an additional match for the union exceeds the cost for the planner, however, because in addition to the increase in vacancy costs $\kappa/\mu'(\theta)$, the union also takes into account that increasing hiring requires reducing wages, thereby giving up some of the surplus it could have appropriated from firms, captured by the term: $-\frac{n}{1-n}\frac{q'(\theta)}{q(\theta)^2}\frac{\kappa}{\mu'(\theta)}$. This additional cost appears also in the Euler equation (20) for the union with commitment, but here it appears both today and tomorrow symmetrically, unlike in the commitment solution where tomorrow's union simply carries out today's plan. Beyond this difference, the union also takes into account its inability to commit to future wages: Creating more matches today will reduce hiring tomorrow, as tomorrow's union will raise wages to exploit those matches. A marginal increase in matches reduces hiring by $\mu'(\theta)(\Theta'(n)(1-n) - \Theta(n))$, with each lost worker valued at the size of the distortion in the union objective—the marginal surplus appropriated from capitalists.

Note that equation (26) differs from standard Euler equations in that the derivative of the function Θ appears in the equation. This means that even solving for a steady state will be more complicated than usual, requiring information about the shape of the Θ function. Steady state refers here to a level of initial matches n and corresponding market tightness $\theta = \Theta(n)$ such that the law of motion maintains the same level of matches: $N(n, \Theta(n)) = n$. In this case, one cannot simply use equation (26) together with the law of motion to solve for a steady state (n, θ)-pair because the derivative appears as an additional unknown.

It is hard to establish theoretically that $\Theta(n)$ is indeed decreasing. In the one-period example of Section 3.1, Θ became a decreasing function of n, and in our numerically solved examples below, this also holds. What is possible to show for the infinite-horizon case, however, is that whenever $\Theta(n)$ is decreasing, steady-state market tightness is strictly below its efficient level.

Proposition 2. *If $\Theta(n)$ is decreasing in n, then the steady-state market tightness, θ, in the unionized labor market (without commitment) is strictly below its efficient level.*

It follows that steady-state unemployment in the unionized labor market is strictly above its efficient level.

3.5. A nonegalitarian union

Relaxing the equal pay constraint by allowing the union to pay different wages to newly hired workers (w_t^n) and workers in existing matches (w_t^e), the union objective becomes

$$\sum_{t=0}^{\infty} \beta^t [n_t w_t^e + \mu(\theta_t)(1-n_t)w_t^n + (1-n_t)(1-\mu(\theta_t))b], \tag{27}$$

and the zero-profit condition

$$\kappa = q(\theta_t)[z - w_t^n + \sum_{s=1}^{\infty} \beta^s (1-\delta)^s (z - w_{t+s}^e)]. \tag{28}$$

In this case a separate condition must be imposed, to ensure that firms make a nonnegative present value of profits on existing workers:

$$\sum_{s=0}^{\infty} \beta^s (1-\delta)^s (z - w_{t+s}^e) \geq 0, \forall t \geq 0. \tag{29}$$

The nonegalitarian union chooses two sequences of wages, $\{w_t^n\}_{t=0}^{\infty}$ and $\{w_t^e\}_{t=0}^{\infty}$, to maximize the objective (27) subject to the law of motion (1), zero-profit conditions (28), and constraints (29) holding at each point in time.

In setting the wages of existing workers, the best the union can do is to set $w_t^e = z$ each period, leaving firms with zero surplus on existing matches. The zero-profit condition then implies that $w_t^n = z - \kappa/q(\theta_t)$, $\forall t \geq 0$. Using this expression to substitute out wages in the union objective, it is easy to see that the union problem becomes identical to the planner problem, thus leading to efficient hiring: $\theta_t = \theta^p$, $\forall t \geq 0$. The solution therefore involves a constant and efficient market tightness over time, as well as constant wages that exhibit a tenure premium: $w_t^n = z - \kappa/q(\theta^p)$ and $w_t^e = z$ $\forall t \geq 0$.

Thus, one can conclude that in the infinite horizon setting as well, the union may be able to attain efficient hiring through a wage tenure premium. A potential concern is that the implied wages of new workers may be quite low—they need to be low enough to allow firms to make the entire present value of profits associated with efficient hiring in the first period of the match. In the presence of a binding lower bound on the wages of junior workers, the union wage policy will still involve a tenure premium, but the market tightness will be distorted down.

In sum, wage solidarity comes at a cost in this economy, suggesting a role for tenure premia in union wages as a means to avoid the resulting distortions in hiring. And yet, the empirical evidence does not point to clearly greater returns to tenure in unionized settings. Is this simply because of the measurement problems involved in the empirical work? Or are the distortions perhaps too insignificant in magnitude to warrant giving up (the benefits underlying) wage solidarity? To shed light on this question, the next section turns to a quantitative illustration looking at the impact of the egalitarian union on labor market outcomes.

4. Quantitative illustration

The presence of an egalitarian union affects the levels and dynamics of wages, unemployment, and output in the economy. This section illustrates these effects, in the context of an extended model.

4.1. Extended model

For added realism, the model is first extended to incorporate partial unionization of the labor market and multiperiod union contracting. To this end, it is assumed that: i) a fraction α of workers are covered by union wages, with a worker's union status fixed over time, while the rest bargain their wages individually, and ii) instead of the union recontracting each period, it recontracts in any given period with probability λ, implying that contracts are expected to last $1/\lambda$ periods.[15]

For the nonunion workers in the labor market, one can write standard Bellman equations, which can then be used to derive the following equation for the match surplus:

$$S_t = z - b + \beta(1-\delta)(1-\mu(\theta_{t+1})\gamma)S_{t+1}. \tag{30}$$

[15] Search is modeled as undirected, an assumption that plays a key role in the discussion in Section 4.4. If the search were fully directed, based on union status, the market would separate into two independent parts: one that follows the full unionization model and one following the standard Mortensen-Pissarides model.

The equation uses the fact that nonunion workers bargain their wages individually, such that the bargaining outcome divides the match surplus according to the workers' bargaining power γ: Workers get γS_t and firms $(1-\gamma)S_t$.[16] Note that the surplus equation (30) depends on the union's actions only through the market tightness.

The firms' zero-profit condition can then be written to reflect the presence of both union and nonunion workers in the labor market as

$$\kappa = q(\theta_t)[\alpha(\frac{z}{1-\beta(1-\delta)} - W_t) + (1-\alpha)(1-\gamma)S_t]. \tag{31}$$

As the right-hand side states, firms expect a present value of profit of $(1-\gamma)S_t$ on the $1-\alpha$ nonunion workers, and a present value of profit of $z/(1-\beta(1-\delta)) - W_t$ on the α union workers. The latter hinges on the expected present value of union wages paid out over the course of an employment relationship: $W_t = \sum_{s=0}^{\infty} \beta^s (1-\delta)^s w_{t+s}$.

One can then think about how union wages $\{w_t\}_{t=0}^{\infty}$ are determined, by returning to the union objective in equation (3). As before, one can rewrite this objective using the zero-profit condition (31), arriving at the expression

$$\sum_{t=0}^{\infty} \beta^t [(n_t + \mu(\theta_t)(1-n_t))z + (1-\mu(\theta_t))(1-n_t)b - \theta_t(1-n_t)\frac{\kappa}{\alpha} + \frac{1-\alpha}{\alpha}(1-\gamma)\mu(\theta_t)(1-n_t)S_t] - \frac{n_0\kappa}{\alpha q(\theta_0)} + \frac{1-\alpha}{\alpha}(1-\gamma)n_0 S_0. \tag{32}$$

Comparing this expression with the corresponding expression (17) earlier, note that with partial unionization, the nonunion surpluses enter into the union objective because of their impact on vacancy creation.

The next step would be to implement multiperiod contracting in this setting, aiming for a recursive representation that could be used to solve the model, as before. Note that as far as union wages are concerned, the object of interest for both the union and the firms is the expected present value of wages paid out over the course of an employment relationship, W_t. This present value determines the profitability of hiring union workers, governing vacancy creation through equation (31). In this sense, the allocative measure of wages here is W_t. What one would like to do, then,

[16] See online appendix, Section C, for a derivation.

is to specify that in periods when the union does not recontract, W_t is held fixed, while in periods when the union does recontract, W_t is reoptimized. With full unionization, this would imply that in periods when the union does not recontract, θ_t remains fixed, while in periods when the union does recontract, θ_t adjusts (due to equation (31)). With partial unionization, this need not hold exactly, because of the presence of the nonunion surpluses in the zero-profit condition. However, it turns out to be clearly simpler to solve the partial union model under the specification that what the union holds fixed in nonrecontracting periods is θ_t directly.[17] This also appears a reasonable approximation to holding W_t fixed, in the sense that changes in W_t during nonrecontracting periods appear minor compared with the adjustments upon recontracting. With these concerns in mind, it is assumed in what follows that what is held fixed in periods when the union does not recontract is θ_t.[18]

To arrive at a recursive representation characterizing labor market outcomes, then, consider first recursive versions of the equations for the nonunion surpluses. Based on equation (30), in periods when the union recontracts, the surplus satisfies:

$$S^r(n) = z - b + \beta(1-\delta)[\lambda(1-\mu(\Theta(N(n,\Theta(n))))\gamma)S^r(N(n,\Theta(n))) + (1-\lambda)(1-\mu(\Theta(n))\gamma)S^f(N(n,\Theta(n)),\Theta(n))], \tag{33}$$

while in periods when the union does not recontract, respectively:

$$S^f(n,\theta) = z - b + \beta(1-\delta)[\lambda(1-\mu(\Theta(N(n,\theta)))\gamma)S^r(N(n,\theta)) + (1-\lambda)(1-\mu(\theta)\gamma)S^f(N(n,\theta),\theta)]. \tag{34}$$

Note that in periods when the union does not recontract, the market tightness is held fixed, while in periods when the union does recontract, the tightness is determined via the equilibrium function $\Theta(n)$. Union decision-making in recontracting periods then determines the function $\Theta(n)$ as the

[17] Solving the partial union model with W_t held fixed leads to systems of nonlinear equations for the nonunion surpluses and their derivatives, while the current specification instead yields linear equations allowing analytical solutions, which is attractive from the point of view of minimizing error associated with numerical complexity.

[18] This distinction matters only for Section 4.4, which allows partial union coverage.

solution to the problem:

$$\Theta(n) \equiv \arg\max_{\theta}(n + \mu(\theta)(1-n))z + (1-\mu(\theta))(1-n)b - \theta(1-n)\frac{\kappa}{\alpha} - \frac{n\kappa}{\alpha q(\theta)} + \frac{1-\alpha}{\alpha}(1-\gamma)(n+\mu(\theta)(1-n))S^r(n) + \beta\lambda V^r(N(n,\theta)) + \beta(1-\lambda)V^f(N(n,\theta),\theta), \quad (35)$$

where the union value satisfies

$$V^r(n) = (n + \mu(\Theta(n))(1-n))z + (1-\mu(\Theta(n)))(1-n)b - \Theta(n)(1-n)\frac{\kappa}{\alpha} + \frac{1-\alpha}{\alpha}(1-\gamma)\mu(\Theta(n))(1-n)S^r(n) + \beta\lambda V^r(N(n,\Theta(n))) + \beta(1-\lambda)V^f(N(n,\Theta(n)),\Theta(n)) \quad (36)$$

in recontracting periods, and

$$V^f(n,\theta) = (n + \mu(\theta)(1-n))z + (1-\mu(\theta))(1-n)b - \theta(1-n)\frac{\kappa}{\alpha} + \frac{1-\alpha}{\alpha}(1-\gamma)\mu(\theta)(1-n)S^f(n,\theta) + \beta\lambda V^r(N(n,\theta)) + \beta(1-\lambda)V^f(N(n,\theta),\theta) \quad (37)$$

in nonrecontracting periods. These equations follow from the union objective (32) as before.

The next section proceeds to calibrating and illustrating the impact of unions in the context of this model. The focus will, for the most part, be on steady states: A level of initial matches n and a corresponding tightness $\theta = \Theta(n)$, such that $N(n,\theta) = n$. With this level of initial matches, if the union recontracts today, it will keep the market tightness unchanged, leading to the same level of initial matches next period.

4.2. Calibration and solution approach

The model is parameterized such that the efficient outcome corresponds to the U.S. labor market, to then study how introducing the union changes outcomes in this market.[19] The period length is set to one month, and the discount rate to correspond to a 5 percent annual rate of return, with $\beta = 1.05^{-12}$. Labor productivity for the market technology is normalized to $z = 1$ and for the home technology set at $b = 0.4$.[20] The matching function is $m(v,u) = \mu_0 vu/(v+u)$, as in den Haan et al.

[19] For consistency, the parameterization strategy follows that described in Shimer (2005), aside from adopting a matching function which is better suited for a discrete time model. He calibrates a decentralized labor market to the U.S. labor market, but the calibration strategy implies that the equilibrium outcome coincides with the socially optimal one.

[20] The results for levels do not change substantially if one raises this to $b = 0.75$.

(2000). The remaining parameters δ, κ, and μ_0, are pinned down as follows: First, attaining an average duration of employment of 2.5 years requires a separation rate of $\delta = 0.033$. Second, to be consistent with a steady-state unemployment rate of 5 percent, the average job-finding rate must be $\mu(\theta) = 0.388$. Finally, to also match the slope of the Beveridge curve, documented by Shimer (2007) to equal -1, this requires setting $\mu_0 = 0.652$ and a steady-state value of $\theta = 1.47$. The latter can be achieved by setting $\kappa = 0.109$.

The basic Mortensen-Pissarides model is straightforward to solve, as is the planner problem previously discussed. The union problem without commitment is clearly more challenging, however. Issues to bear in mind include the fact that there are few results on the existence of equilibrium for differentiable Markov perfect equilibria; that these equilibria may not be unique and that nondifferentiable equilibria may exist as well.[21] In solving for a differentiable equilibrium, a natural starting point would be the generalized Euler equation of the problem. Here, the complexity of the system (33–37) does not allow us to derive such an equation explicitly, but it turns out that one can proceed along the same lines without this formal step. The focus will be on steady states; the solution method adopted follows the approach of Krusell et al. (2002), which looks for a Taylor expansion approximation to the unknown function $\Theta(n)$ around the steady state. The approach involves solving successively larger systems of equations based on the first-order condition (and successive derivatives of the first-order condition) of problem (35), looking for convergence in the coefficients of the polynomial as the order increases. A description of how the approach is implemented here can be found in the online appendix, Section D. The next sections describe the results.

4.3. Level effects

This section begins by looking at the impact that introducing the union has on the levels of wages, unemployment, and output, relative to the efficient outcome, in the case of full coverage. To that end, recall from the theory that the duration of union contracts should be an important de-

[21] For examples where no differentiable equilibria exist but a nondifferentiable equilibrium does, see Krusell et al. (2005), and for examples with a continuum of nondifferentiable equilibria along with one or more differentiable ones, see Krusell and Smith (2003); Phelps and Pollak (1968) focus on differentiable equilibria and find multiplicity as well.

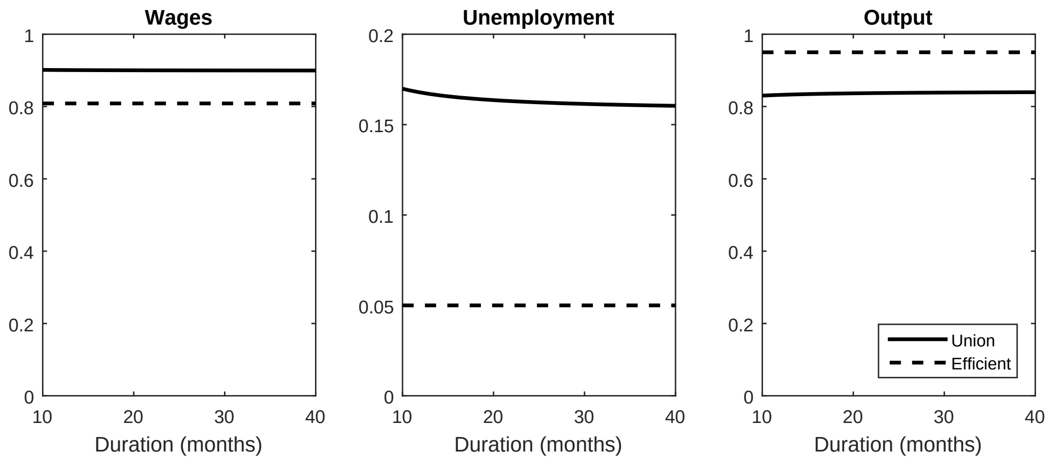


Figure 2: Level effect of union

Notes: The figure plots steady-state wages, unemployment, and output, as a function of the expected duration of a union contract $1/\lambda$.

terminant of the magnitude of the distortions associated with the union. Available evidence seems to point to one to three years as the relevant range of union contract durations, and accordingly λ is set to $1/24$, implying an expected duration of union contracts of two years.[22] With this duration of union contracts, introducing the union into the labor market is found to raise wages by 11 percent, leading to an increase in unemployment from 5 to 16 percent, and a reduction in output of 12 percent, relative to efficient outcomes. As expected, wages and unemployment thus rise, leading to lower output, but the calculation reveals the quantitative impact to be substantial as well.

To see how the effects depend on contract duration, Figure 2 plots the steady-state levels of wages, unemployment, and output as a function of the expected duration, $1/\lambda$. The benchmark in the figure—the efficient outcome—is naturally independent of λ. The figure shows that the impact of the union diminishes as contract duration increases, as one would expect. But the figure also reveals that for the relevant range of contract durations this effect turns out to be rather weak. Even though there is a visible decrease in unemployment as contract duration increases from one to three

[22]For example, for the U.S., Taylor (1983) considers one to three years as the relevant range of union contracts, Card (1990) documents an average contract duration of 26 months, and Rich and Tracy (2004) a median duration of 36 months. Fregert and Jonung (2006) document similar durations for Sweden, and Avouyi-Dovi et al. (2013) an average duration of just under a year for France.

years, the magnitude of this decrease is overshadowed by the overall level effect associated with the union. Note that there is no reason to expect the union outcome to converge with the efficient one as the duration of contracts approaches infinity: Recall that in the commitment union problem analyzed in Section 3.3, the union distorts θ down in the initial period but attains the efficient θ thereafter. This multiperiod contracting specification, on the other hand, constrains θ to remain fixed between recontracting periods. Thus, it would seem natural for the union to set this fixed tightness above the efficient level when recontracting.

Finally, recall that the decentralized outcome in the Mortensen-Pissarides model is efficient only if the private bargaining power of workers coincides with the one implementing efficient allocations (Hosios, 1990). Unemployment can thus exceed the efficient level also in the decentralized equilibrium, if workers are strong bargainers. The next section returns to this issue, in considering the case of partial union coverage.

4.4. Union coverage

In a classic paper, Calmfors and Driffill (1988) reconsider the impact of unions on the level of economic activity. It has long been recognized that unions, through their monopoly power in the labor market, tend to raise wages above their competitive levels. This suggests that a greater union presence in the labor market has a primarily negative impact on economic activity, as high union wages lead to higher unemployment. Calmfors and Driffill (1988) propose an additional factor for understanding the cross-country evidence on unions: They argue that the degree of coordination in union bargaining works to counteract the negative effects of monopoly power. A related hump-shaped relationship emerges in our model as well, when the coverage of union wages across the workforce is varied.

Two competing forces come to play in the model as union coverage varies: First of all, because union wages tend to exceed nonunion wages, greater union coverage tends to lead to higher unemployment here as well. But greater union coverage also increases the extent to which the union takes into account the effects of its wage demands on hiring, borne by union and nonunion workers alike, leading to moderation in union wage setting. As union coverage increases, the second effect eventually takes over the first, leading to a hump-shaped relationship between union coverage and

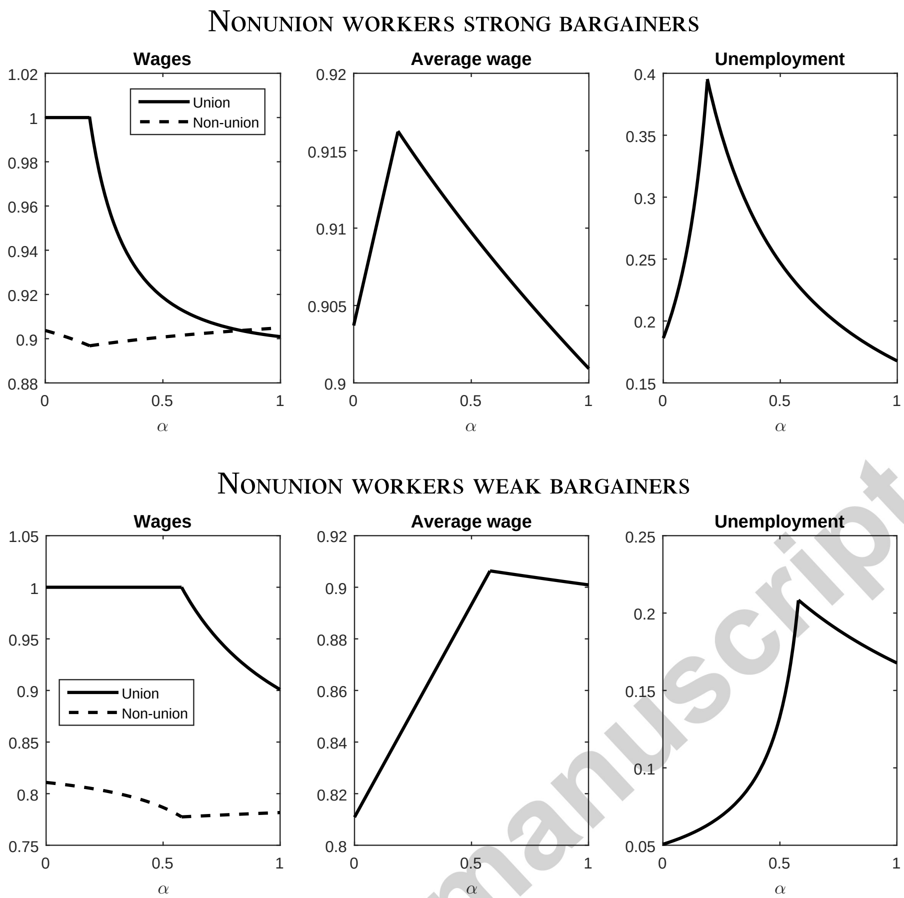


Figure 3: Role of unionization rate

Notes: The figure plots union and nonunion wages, the average wage, and unemployment as a function of union coverage α. The nonunion bargaining power is $\gamma = 0.8$ in the top panel, and $\gamma = 0.6$ in the bottom panel.

unemployment.

Figure 3 illustrates the relationship between union coverage and unemployment in the model, contrasting two cases that differ in the bargaining power of the nonunion workers in their private wage bargains. In the first, nonunion workers are strong bargainers, with $\gamma = 0.8$. The top panels of Figure 3 plot the steady-state levels of wages and unemployment in this case. The plot on the left first shows how union and nonunion wages vary with union coverage. As union coverage falls, union wages rise until they equal productivity and cannot rise further. In the meantime, the wages of nonunion workers remain mostly unaffected, although they reflect changes in the outside options of these workers, which are worse at intermediate levels of coverage. What enters into firms' profits is the weighted average of these wages across the pool of unemployed shown in the middle. Averaging across workers yields a hump-shaped relationship between union coverage and the average wage, which further gives rise to the hump-shaped relationship between union coverage and unemployment shown on the right.

Note that unemployment well exceeds the efficient level of 5 percent here even without the union because of the high private bargaining power of workers, and that introducing the union can improve outcomes over that alternative, if the coverage is high enough. One could also ask what level of union coverage would be expected to emerge if workers could choose (in the beginning of time) whether to be union or nonunion. In Figure 3, an interior union coverage level exists where workers would be indifferent between being union versus nonunion in terms of the wages being equal. At that coverage level, unemployment is lower than it would be if unions were outlawed completely but higher than with universal coverage of union wages.

To see how the picture changes when workers are weaker bargainers, the bottom panels of Figure 3 consider the case in which the worker's bargaining power yields efficient outcomes (here $\gamma = 0.6$). The figure is qualitatively similar, but in this case, unemployment is always higher in the unionized labor market than it would be without the union. Union wages also always exceed nonunion wages and by a clear margin. Given a choice, all workers would prefer to be in the union, but it would be welfare improving to outlaw the union instead.

4.5. Shock propagation

An important reason that macroeconomists have been interested in labor unions is the notion that unions create rigidity in wages, affecting how the economy responds to shocks (see, e.g., Blanchard and Fischer, 1989, pp. 438–455). This section illustrates the impact of unions on shock propagation, in the context of our model with full unionization.

The focus is on the effects of a one-time, unanticipated, permanent increase in labor productivity. Having first solved for the steady state before the shock, one can then look at how the transition to higher productivity plays out when the expected duration of wage contracts is two years. Figure 4 plots the responses, comparing the unionized labor market (solid line) with the efficient (dashed line), as well as fully fixed wages (dotted line). In the efficient response, the

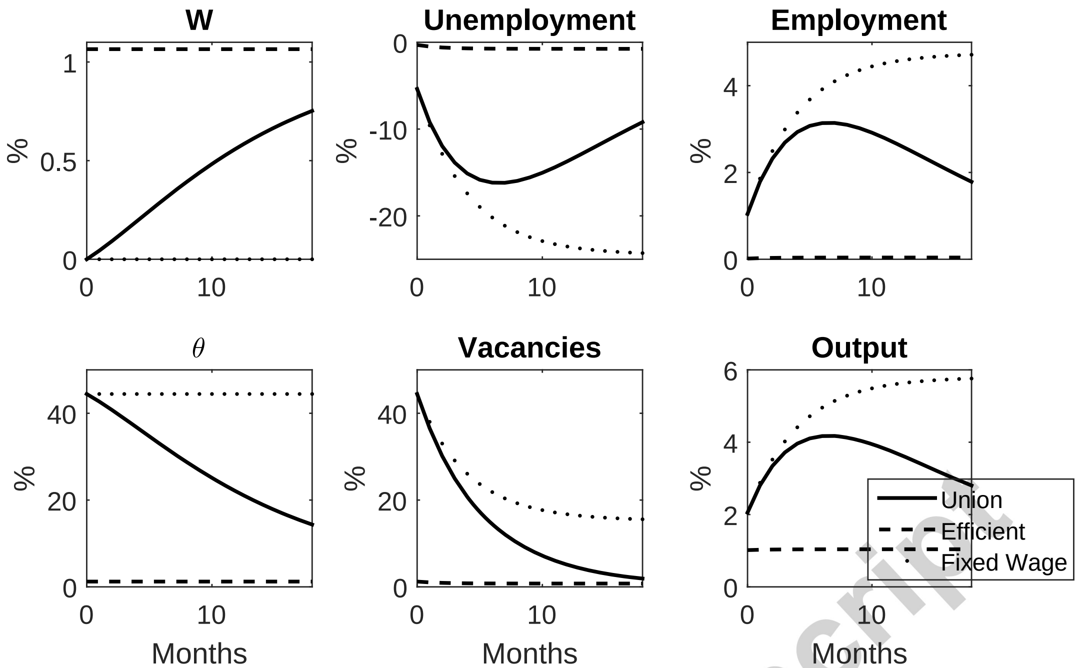


Figure 4: Impulse responses

Notes: The figure plots the responses of the present value of wages, market tightness, unemployment, vacancies, employment, and output to a 1 percent unanticipated permanent increase in productivity. The figure shows the response for the economy with full coverage of union wages with two-year contracts, the efficient response, and the response with fully fixed wages. What is plotted are expected values in each period after the increase in productivity, across possible realizations of the recontracting shock.

wage and market tightness adjust immediately to their new steady-state levels. With fixed wages, the market tightness also adjusts immediately to its new steady-state level, although in this case, larger than what is efficient. The union response lies between these two extremes but also differs in exhibiting significant inertia in wages because of the multiperiod union contracting.

In terms of the magnitudes of these responses, the efficient response reflects a sizable on-impact response of the wage to the shock, which leads to small responses in quantities. This is the unemployment volatility puzzle discussed by Shimer (2005): The magnitude of these responses is an order of magnitude lower than what would be needed to be consistent with business cycle fluctuations in the data. If wages are fixed in response to the shock, quantities respond substantially more strongly, as highlighted by Hall (2005), allowing the model to match the magnitude of fluctuations observed. The stickiness in union wages, with two-year contracting, increases the volatility of quantities substantially relative to the efficient responses.

Hidden behind the stickiness in wages associated with multiperiod contracting, there is also a mechanism generating endogenous real wage rigidity in the model: The wage increase in response to the increase in productivity takes some time to play out, as the union distortion which works to raise wages becomes stronger as matches accumulate over time after the shock. In Figure 4, the quantitative impact of this endogenous rigidity is overwhelmed by that of the stickiness associated with multiperiod contracting, however.

5. Conclusions

A holdup problem emerges when an egalitarian union sets wages in a frictional labor market. After demonstrating the issue in a theoretical setting, this paper studies the severity of the holdup problem quantitatively in an extended model with partial union coverage and multiperiod union contracting. It is shown to raise wages and unemployment significantly above their efficient levels. The relationship between union coverage and unemployment is hump-shaped in the model, with intermediate levels of coverage featuring higher unemployment than either very low or very high coverage, and the bargaining power of nonunion workers playing a key role in determining which of the two extremes is closer to efficient allocations. Multi-period union contracts generate significant stickiness in the response of wages to shocks. Finally, the theory implies a rationale for a tenure premium in union wages, as a means of avoiding the distortions associated with holdup.

The analysis is conducted in a stylized setting, to isolate key forces at play, but many extensions would seem natural, such as incorporating market power/decreasing returns, physical capital, worker heterogeneity, an insider-outsider wedge, as well as thinking more about the decisions of workers to join versus leave the union in a dynamic setting.

Abraham, K. G., Farber, H. S., 1988. Returns to seniority in union and nonunion jobs: a new look at the evidence. Industrial and Labor Relations Review 42 (1), 3–19.

Acemoglu, D., Pischke, J.-S., 1999. The structure of wages and investment in general training. Journal of Political Economy 107, 539–572.

Acikgoz, O., Kaymak, B., 2014. The rising skill premium and deunionization in the United States. Journal of Monetary Economics 63, 37–50.

Altonji, J., Williams, N., 2005. Do wages rise with job seniority? a reassessment. Industrial and Labor Relations Review 58, 370–97.

Alvarez, F., Shimer, R., 2011. Unions and unemployment, unpublished manuscript, University of Chicago.

Alvarez, F., Veracierto, M., 2000. Labor-market policies in an equilibrium search model. NBER Macroeconomics Annual, 265–316.

Avouyi-Dovi, S., Fougere, D., Gautier, E., 2013. Wage rigidity, collective bargaining, and the minimum wage: Evidence from french agreement data. Review of Economics and Statistics 94 (4), 1337–51.

Blanchard, O., Fischer, S., 1989. Lectures on Macroeconomics. The MIT Press, Cambridge, Massachusetts.

Blanchflower, D. G., Bryson, A., 2003. Changes over time in union relative wage effects in the U.K. and the U.S. revisited. In: Addison, J. T., Schnabel, C. (Eds.), International Handbook of Trade Unions. Edward Elgar Publishing, Ch. 7, pp. 197–245.

Boeri, T., Burda, M. C., 2009. Preferences for collective versus individualised wage setting. Economic Journal 119, 1440–1463.

Buchinsky, M., Fougre, D., Kramarz, F., Tchernis, R., 2010. Interfirm mobility, wages and the returns to seniority and experience in the united states. Review of Economic Studies 77 (3), 972–1001.

Calmfors, L., Driffill, J., 1988. Bargaining structure, corporatism and macroeconomic performance. Economic Policy 88, 13–61.

Card, D., 1990. Unexpected inflation, real wages, and employment determination in union contracts. American Economic Review 80 (4), 669–88.

Card, D., Lemieux, T., Riddell, C., 2003. Unions and the wage structure. In: Addison, J. T., Schnabel, C. (Eds.), International Handbook of Trade Unions. Edward Elgar Publishing, Ch. 8, pp. 246–292.

Delacroix, A., 2006. A multisectorial matching model of unions. Journal of Monetary Economics 53, 573–596.

den Haan, W. J., Ramey, G., Watson, J., 2000. Job destruction and propagation of shocks. American Economic Review 90 (3), 482–498.

DiNardo, J., Lee, D. S., 2004. Economic impacts of new unionization on private sector employers: 1984–2001. Quarterly Journal of Economics 119 (4), 1383–1441.

Ebell, M., Haefke, C., 2006. Product market regulation and endogenous union formation, IZA Discussion Paper No. 2222.

Fregert, K., Jonung, L., 2006. Policy rule evaluation by contract-makers: 100 years of wage contract length in Sweden, European Comission Economic Papers.

Garibaldi, P., Violante, G., 2005. Employment effects of severance payments with wage rigidities. Economic Journal 115, 799–832.

Hall, R. E., 2005. Employment fluctuations with equilibrium wage stickiness. American Economic Review 95 (1), 50–65.

Hosios, A., 1990. On the efficiency of matching and related models of search and unemployment. Review of Economic Studies 57, 279–298.

Krusell, P., Kuruscu, B., Smith, A., 2002. Equilibrium welfare and government policy with quasi-geometric discounting. Journal of Economic Theory 105, 42–72.

Krusell, P., Martin, F., Rios-Rull, V., 2005. Time consistent debt, unpublished manuscript, University of Rochester.

Krusell, P., Smith, A., 2003. Consumption-savings decisions with quasi-geometric discounting. Econometrica 71, 365–375.

Lewis, H. G., 1986. Union relative wage effects. In: Ashenfelter, O., Layard, R. (Eds.), Handbook of Labor Economics. Elsevier, Ch. 20, pp. 1140–81.

Lockwood, B., Manning, A., 1989. Wage-employment bargaining with employment adjustment costs. Economic Journal 99, 1143–58.

Lucas, Jr, R., Prescott, E. C., 1974. Equilibrium search and unemployment. Journal of Economic Theory 7, 188–209.

Modesto, L., Thomas, J. P., 2001. An analysis of labor adjustment costs in unionized economies. Labour Economics 8, 475–501.

Nickell, S., Layard, R., 1999. Labor market institutions and economic performance. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics. Vol. 3. Elsevier, Ch. 46, pp. 3029–84.

Phelps, E., Pollak, R., 1968. On second-best national saving and game-equilibrium growth. Review of Economic Studies 35, 185–199.

Pissarides, C., 1985. Short-run equilibrium dynamics of unemployment, vacancies, and real wages. American Economic Review 75 (4), 676–690.

Pissarides, C., 1986. Trade unions and the efficiency of the natural rate of unemployment. Journal of Labor Economics 4 (4), 582–595.

Rich, R., Tracy, J., 2004. Uncertainty and labor contract durations. Review of Economics and Statistics 86 (1), 270–87.

Shimer, R., 2005. The cyclical behavior of equilibrium unemployment and vacancies. American Economic Review 95 (1), 25–49.

Shimer, R., 2007. Mismatch. American Economic Review 97 (4), 1074–1101.

Taschereau-Dumouchel, M., 2011. The union threat, unpublished manuscript, Princeton University.

Taylor, J. B., 1983. Union wage settlements during a disinflation. American Economic Review 73 (5), 981–993.

Topel, R., 1991. Specific capital, mobility, and wages. Journal of Political Economy 99 (1), 145–176.

Visser, J., 2003. Unions and unionism around the world. In: Addison, J. T., Schnabel, C. (Eds.), International Handbook of Trade Unions. Edward Elgar Publishing, Ch. 11, pp. 366–413.

Visser, J., Hayter, S., Gammarano, R., 2015. Trends in collective bargaining coverage: stability, erosion or decline?, ILO: Labour Relations and Collective Bargaining Issues Brief No. 1.